

EN998028

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2

INTERNATIONAL BUSINESS MACHINES CORPORATION

**ADAPTIVELY ENCODING A PICTURE OF
CONTRASTED COMPLEXITY HAVING NORMAL
VIDEO AND NOISY VIDEO PORTIONS**

Technical Field

5 This invention relates in general to compression
of digital visual images, and more particularly, to a
technique for encoding one or more frames of
contrasted complexity within a video sequence using
image statistics derived from the frame(s) to
10 dynamically change one or more controllable encoding
parameter(s) used in encoding the frame(s).

Background of the Invention

15 Within the past decade, the advent of world-wide
electronic communications systems has enhanced the
way in which people can send and receive information.
In particular, the capabilities of real-time video
and audio systems have greatly improved in recent
years. However, in order to provide services such as
video-on-demand and video conferencing to
20 subscribers, an enormous amount of network bandwidth
is required. In fact, network bandwidth is often the
main inhibitor in the effectiveness of such systems.

25 In order to overcome the constraints imposed by
networks, compression systems have emerged. These
systems reduce the amount of video and audio data
which must be transmitted by removing redundancy in
the picture sequence. At the receiving end, the
picture sequence is uncompressed and may be displayed
in real-time.

One example of a video compression standard is the Moving Picture Experts Group ("MPEG") standard. Within the MPEG standard, video compression is defined both within a given picture and between
5 pictures. Video compression within a picture is accomplished by conversion of the digital image from the time domain to the frequency domain by a discrete cosine transform, quantization, and variable length coding. Video compression between pictures is
10 accomplished via a process referred to as motion estimation and compensation, in which a motion vector plus difference data is used to describe the translation of a set of picture elements (pels) from one picture to another.

15 The ISO MPEG-2 standard specifies only the syntax of bitstream and semantics of the decoding process. The choice of coding parameters and trade-offs in performance versus complexity are left to the encoder developers.

20 One aspect of the encoding process is compressing a digital video image into as small a bitstream as possible while still maintaining video detail and quality. The MPEG standard places limitations on the size of the bitstream, and
25 requires that the encoder be able to perform the encoding process. Thus, simply optimizing the bit rate to maintain desired picture quality and detail can be difficult.

A video picture typically contains both busy and
30 simple macroblock segments, and there is a high correlation between the segments. However, certain

video frames are of highly contrasted complexity having, e.g., both normal video and noisy (or random) video portions within the frame, such as DIVA. Further, both the normal (or simple) video portion
5 and the noisy portion are often moving from frame to frame. Within such a frame, most of the encode bits can be consumed by macroblocks of the noisy portion before picture coding is completed, thereby producing blockiness or artifacts within the picture and uneven
10 output picture quality.

This invention thus seeks to enhance picture quality of an encoded video sequence having one or more pictures with areas of significantly contrasted complexity, and more particularly, to enhance picture
15 quality by dynamically balancing picture bit allocation as the picture coding continues without requiring lengthy buffering or high computational intelligence.

Disclosure of the Invention

20 Briefly summarized, the invention comprises in a first aspect a method for encoding a video frame having a noisy portion and a normal video portion. The method includes for each macroblock of the frame: determining a macroblock activity level; determining
25 whether the macroblock activity level exceeds a predefined threshold, wherein the macroblock activity level exceeding the predefined threshold indicates that the macroblock is associated with the noisy portion of the video frame; and adjusting encoding of
30 the macroblock when the macroblock activity level exceeds the threshold to conserve bits used in

encoding the macroblock and thereby reduce the number of bits used to encode macroblocks within the noisy portion of the video frame.

5 In another aspect, a method is presented for encoding a frame of a sequence of video frames, each frame having a plurality of macroblocks. The method includes: determining whether the frame includes a random noise portion; and when the frame does include a random noise portion, evaluating each macroblock of
10 the plurality of macroblocks in the frame and adjusting encoding of at least some macroblocks within the random noise portion of the frame, the adjusting of encoding comprising conserving bits used in encoding the at least some macroblocks within the
15 random noise portion of the frame.

In still another aspect, a system for encoding a frame having a noisy portion is provided. The system includes means for determining a macroblock activity level and means for determining when the macroblock
20 activity level exceeds a predefined threshold. The macroblock activity level exceeding the predefined threshold is indicative that the macroblock is associated with the noisy portion of the frame. The system further includes means for adjusting encoding
25 of the macroblock when the macroblock activity level exceeds the predefined threshold in order to reduce bits used in encoding the macroblock, and thereby conserve bits otherwise used to encode macroblocks within the noisy portion of the frame.

30 In a further aspect, a system is provided for encoding a frame of a sequence of frames. This

system includes a pre-encode processing unit for determining whether the frame includes a random noise portion, and a control and encode unit for evaluating each macroblock of a plurality of macroblocks comprising the frame when the frame includes the random noise portion. The control and encode unit includes means for adjusting encoding of at least some macroblocks within the random noise portion of the frame to reduce bits used in encoding the macroblocks within the random noise portion.

In still other aspects, the concepts presented herein are implemented within computer program products having computer usable medium with computer readable program code means therein for use in encoding a frame as summarized above.

Advantageously, processing in accordance with the present invention prevents noisy macroblocks or blocks with random details from consuming all or most of the picture bits, which in turn prevents overproduction of bits before the encoder reaches the bottom of the picture. This invention essentially directs encode bits from the random, busy macroblocks to the simpler, normal macroblocks. Less bits are used in the highly active and fine detailed area, thereby providing a more constant picture quality.

Brief Description of the Drawings

The above-described objects, advantages and features of the present invention, as well as others, will be more readily understood from the following detailed description of certain preferred embodiments

of the invention, when considered in conjunction with the accompanying drawings in which:

Fig. 1 shows a flow diagram of a generalized MPEG-2 compliant encoder 11, including a discrete cosine transformer 21, a quantizer 23, a variable length coder 25, an inverse quantizer 29, an inverse discrete cosine transformer 31, motion compensation 41, frame memory 42, and motion estimation 43. The data paths include the i^{th} picture input 111, difference data 112, motion vectors 113 (to motion compensation 41 and to variable length coder 25), the picture output 121, the feedback picture for motion estimation and compensation 131, and the motion compensated picture 101. This figure has the assumptions that the i^{th} picture exists in frame memory or frame store 42 and that the $i+1^{\text{th}}$ is being encoded with motion estimation.

Fig. 2 illustrates the I, P, and B pictures, examples of their display and transmission orders, and forward, and backward motion prediction.

Fig. 3 illustrates the search from the motion estimation block in the current frame or picture to the best matching block in a subsequent or previous frame or picture. Elements 211 and 211' represent the same location in both pictures.

Fig. 4 illustrates the movement of blocks in accordance with the motion vectors from their position in a previous picture to a new picture, and the previous picture's blocks adjusted after using motion vectors.

Fig. 5 depicts one embodiment of a frame of contrasted complexity having normal video and noisy random video portions to be processed in accordance with the adaptive encoding of the present invention.

5 **Fig. 6** shows a generalized encode system 300 in accordance with the present invention. System 300 includes pre-encode statistics analysis 310 to determine whether an input picture comprises a picture of contrasted complexity and based thereon
10 whether one or more encoding parameters should be varied for individual macroblocks of the picture. The modified encoding parameters are used by encode engine 320 in encoding the individual macroblocks of the picture.

15 **Fig. 7** is a flowchart of one embodiment of identifying a current frame of a sequence of video frames as comprising a frame with a noisy or random portion for processing in accordance with the present invention.

20 **Fig. 8** is a flowchart of one embodiment of adaptively encoding a picture having a noisy video portion in accordance with the present invention.

Best Mode for Carrying Out the Invention

25 The invention relates, for example, to MPEG compliant encoders and encoding processes such as described in "Information Technology-Generic coding of moving pictures and associated audio information: Video," Recommendation ITU-T H.262, ISO/IEC 13818-2, Draft International Standard, 1994. The encoding

functions performed by the encoder include data
input, spatial compression, motion estimation,
macroblock type generation, data reconstruction,
entropy coding, and data output. Spatial compression
5 includes discrete cosine transformation (DCT),
quantization, and entropy encoding. Temporal
compression includes intensive reconstructive
processing, such as inverse discrete cosine
transformation, inverse quantization, and motion
10 compensation. Motion estimation and compensation are
used for temporal compression functions. Spatial and
temporal compression are repetitive functions with
high computational requirements.

Further, the invention relates, for example, to
15 a process for performing spatial and temporal
compression including discrete cosine transformation,
quantization, entropy encoding, motion estimation,
motion compensation, and prediction, and even more
particularly to a system for accomplishing spatial
20 and temporal compression.

The first compression step is the elimination of
spatial redundancy, for example, the elimination of
spatial redundancy in a still picture of an "I" frame
picture. Spatial redundancy is the redundancy within
25 a picture. The MPEG-2 Draft Standard is using a
block based method of reducing spatial redundancy.
The method of choice is the discrete cosine
transformation, and discrete cosine transform coding
of the picture. Discrete cosine transform coding is
30 combined with weighted scalar quantization and run
length coding to achieve desirable compression.

The discrete cosine transformation is an orthogonal transformation. Orthogonal transformations, because they have a frequency domain interpretation, are filter bank oriented. The
5 discrete cosine transformation is also localized. That is, the encoding process samples on an 8x8 spatial window which is sufficient to compute 64 transform coefficients or sub-bands.

Another advantage of the discrete cosine
10 transformation is that fast encoding and decoding algorithms are available. Additionally, the sub-band decomposition of the discrete cosine transformation is sufficiently well behaved to allow effective use of psychovisual criteria.

15 After transformation, many of the frequency coefficients are zero, especially the coefficients for high spatial frequencies. These coefficients are organized into a zig-zag or alternate-scanned pattern, and converted into run-amplitude (run-level)
20 pairs. Each pair indicates the number of zero coefficients and the amplitude of the non-zero coefficient. This is coded in a variable length code.

Motion compensation is used to reduce or even
25 eliminate redundancy between pictures. Motion compensation exploits temporal redundancy by dividing the current picture into blocks, for example, macroblocks, and then searching in previously transmitted pictures for a nearby block with similar
30 content. Only the difference between the current block pels and the predicted block pels extracted

from the reference picture is actually compressed for transmission and thereafter transmitted.

The simplest method of motion compensation and prediction is to record the luminance and chrominance, i.e., intensity and color, of every pixel in an "I" picture, then record changes of luminance and chrominance, i.e., intensity and color for every specific pixel in the subsequent picture. However, this is uneconomical in transmission medium bandwidth, memory, processor capacity, and processing time because objects move between pictures, that is, pixel contents move from one location in one picture to a different location in a subsequent picture. A more advanced idea is to use a previous or subsequent picture to predict where a block of pixels will be in a subsequent or previous picture or pictures, for example, with motion vectors, and to write the result as "predicted pictures" or "P" pictures. More particularly, this involves making a best estimate or prediction of where the pixels or macroblocks of pixels of the i^{th} picture will be in the $i-1^{\text{th}}$ or $i+1^{\text{th}}$ picture. It is one step further to use both subsequent and previous pictures to predict where a block of pixels will be in an intermediate or "B" picture.

To be noted is that the picture encoding order and the picture transmission order do not necessarily match the picture display order. See **Fig. 2**. For I-P-B systems the input picture transmission order is different from the encoding order, and the input pictures must be temporarily stored until used for

encoding. A buffer stores this input until it is used.

For purposes of illustration, a generalized flowchart of MPEG compliant encoding is shown in **Fig. 1**. In the flowchart the images of the i^{th} picture and the $i+1^{\text{th}}$ picture are processed to generate motion vectors. The motion vectors predict where a macroblock of pixels will be in a prior and/or subsequent picture. The use of the motion vectors is a key aspect of temporal compression in the MPEG standard. As shown in **Fig. 1** the motion vectors, once generated, are used for the translation of the macroblocks of pixels, from the i^{th} picture to the $i+1^{\text{th}}$ picture.

As shown in **Fig. 1**, in the encoding process, the images of the i^{th} picture and the $i+1^{\text{th}}$ picture are processed in the encoder 11 to generate motion vectors which are the form in which, for example, the $i+1^{\text{th}}$ and subsequent pictures are encoded and transmitted. An input image 111 of a subsequent picture goes to the motion estimation unit 43 of the encoder. Motion vectors 113 are formed as the output of the motion estimation unit 43. These vectors are used by the motion compensation Unit 41 to retrieve macroblock data from previous and/or future pictures, referred to as "reference" data, for output by this unit. One output of the motion compensation Unit 41 is negatively summed with the output from the motion estimation unit 43 and goes to the input of the Discrete Cosine Transformer 21. The output of the discrete cosine transformer 21 is quantized in a quantizer 23. The output of the quantizer 23 is split

into two outputs, 121 and 131; one output 121 goes to a downstream element 25 for further compression and processing before transmission, such as to a run length encoder; the other output 131 goes through reconstruction of the encoded macroblock of pixels for storage in frame memory 42. In the encoder shown for purposes of illustration, this second output 131 goes through an inverse quantization 29 and an inverse discrete cosine transform 31 to return a lossy version of the difference macroblock. This data is summed with the output of the motion compensation unit 41 and returns a lossy version of the original picture to the frame memory 42.

As shown in **Fig. 2**, there are three types of pictures. There are "Intra pictures" or "I" pictures which are encoded and transmitted whole, and do not require motion vectors to be defined. These "I" pictures serve as a reference image for motion estimation. There are "Predicted pictures" or "P" pictures which are formed by motion vectors from a previous picture and can serve as a reference image for motion estimation for further pictures. Finally, there are "Bidirectional pictures" or "B" pictures which are formed using motion vectors from two other pictures, one past and one future, and can not serve as a reference image for motion estimation. Motion vectors are generated from "I" and "P" pictures, and are used to form "P" and "B" pictures.

One method by which motion estimation is carried out, shown in **Fig. 3**, is by a search from a macroblock 211 of an i^{th} picture throughout a region of the next picture to find the best match macroblock

213. Translating the macroblocks in this way yields
a pattern of macroblocks for the $i+1^{\text{th}}$ picture, as
shown in **Fig. 4**. In this way the i^{th} picture is
changed a small amount, e.g., by motion vectors and
5 difference data, to generate the $i+1^{\text{th}}$ picture. What
is encoded are the motion vectors and difference
data, and not the $i+1^{\text{th}}$ picture itself. Motion vectors
translate position of an image from picture to
picture, while difference data carries changes in
10 chrominance, luminance, and saturation, that is,
changes in shading and illumination.

Returning to **Fig. 3**, we look for a good match by
starting from the same location in the i^{th} picture as
in the $i+1^{\text{th}}$ picture. A search window is created in
15 the i^{th} picture. We search for a best match within
this search window. Once found, the best match motion
vectors for the macroblock are coded. The coding of
the best match macroblock includes a motion vector,
that is, how many pixels in the y direction and how
20 many pixels in the x direction is the best match
displaced in the next picture. Also encoded is
difference data, also referred to as the "prediction
error", which is the difference in chrominance and
luminance between the current macroblock and the best
25 match reference macroblock.

The operational functions of an MPEG-2 encoder
are discussed in detail in commonly assigned, co-
pending United States Patent Application Serial No.
08/831,157, by Carr et al., filed April 1, 1997,
30 entitled "Control Scheme For Shared-Use Dual-Port
Predicted Error Array," which is hereby incorporated
herein by reference in its entirety.

Encoder performance and picture quality are often enhanced today through the use of adaptive quantization. Examples of adaptive quantization are presented in co-pending, commonly assigned United States Patent Applications by Boroczky et al.,
5 entitled "Adaptive Real-Time Encoding of Video Sequence Employing Image Statistics," filed October 10, 1997, serial no. 08/948,442, and by Boice et al.,
10 entitled "Real-Time Variable Bit Rate Encoding of Video Sequence Employing Image Statistics," filed January 16, 1998, serial no. 09/008,282, both of which are hereby incorporated herein by reference in their entirety.

Adaptive quantization can be used to control the amount of data generated so that an average amount of data is output by the encoder and so that this average will match a specified bitrate. As one approach, video quality of a picture having a noisy video portion can be balanced by channeling bits from
15 the noisy or high activity macroblocks to the normal portion of the picture. For example, sophisticated pre-processing might initially be used to determine how picture target bits are to be allocated among all the macroblocks of a picture having noisy video.
20 However, there are 1350 macroblocks in a NTSC picture and 1440 macroblocks in a PAL picture, and the amount of preprocessing logic to accomplish this approach would require significant buffering and a large amount of computational intelligence.

30 As a preferred approach, presented herein is a novel design for dynamically balancing picture bit allocation within a highly contrasted picture having

normal video and noisy video sections as picture
coding continues without significant buffering of the
picture and without requiring large computational
intelligence to accomplish balancing of the bit
5 allocation.

Fig. 5 depicts one embodiment of a picture 250
of contrasted complexity having a random noise
portion 260 and a normal video portion 270. As used
in this application, a "contrasted picture" or
10 "picture of contrasted complexity" means any picture
having a first area of high or random activity and a
second area of significantly lower activity. "Noisy
video" is used herein to denote a picture or that
portion of a picture having very high complexity,
15 such as a picture portion having randomly moving dots
of different color. "Normal video" is used to mean a
picture or portion of a picture depicting, for
example, a conventional motion picture image. **Fig. 5**
is thus shown by way of example only and those
20 skilled in the art will understand that a frame
having contrasted complexity sections of "normal
video" and "noisy video" can encompass many
variations.

In accordance with this invention, the
25 complexity of each input picture is statistically
calculated as the picture is received by the encoder.
This complexity measurement is tailored to indicate
the degree of business or amount of detail within the
picture. From picture complexity, an average
30 complexity value for each macroblock can be
determined. During the macroblock coding process,
the encoder calculates the actual macroblock

complexity and alters the coding options in accordance with this invention when picture complexity is above a predefined, experimentally determined complexity threshold, and the specified
5 bitrate is lower than a predefined bitrate threshold. The complexity and bitrate thresholds can be selected experimentally by one skilled in the art in order to accomplish the objects of the present invention. Basically, this invention seeks to dynamically modify
10 the coding algorithm when the bitrate is too low for the material to be encoded given that the current picture has been statistically determined to comprise a picture having a noisy portion of very high activity.

15 Changes to the coding algorithm can include adjusting the macroblock coding type and modifying the quantization level. For example, once a contrasted picture is identified, the macroblock coding type is preferably biased towards being coded
20 predictive, that is, it requires a larger prediction error before a macroblock will be coded as intra. When the macroblock is coded as intra, the macroblock is thus truly different from the prior reference picture. Since intra macroblocks take many more bits
25 to code than predictive macroblocks, the quantization level of these macroblocks is also adjusted to conserve bits.

For example, a more precise quantization level can be determined from an activity value that is a
30 better representation of the macroblock to be encoded. The relative activity of each block in a macroblock is examined, and the block activity that

is exceptionally far from the rest is discarded. In one embodiment, the block activities can be prioritized and the smallest activity value is compared to the next smallest one. If the block with the smallest amount of activity is one-half or less the block with the next smallest activity, and is one-half or less the average activity within the macroblock, then that block with the lowest activity is preferably ignored in the quantization level calculation. The calculated quantization level can also be increased by a percentage determined from experiments. Again, the goal is to conserve bits when encoding macroblocks of the noisy video portion, thereby providing more bits for encoding macroblocks within the normal video portion.

Fig. 6 depicts one embodiment of an encode system, generally denoted 300, in accordance with this invention. As shown, an input stream of video frames is conventionally buffered in frame memory 330. Controller 340 determines where a given input picture should be placed within the memory, as well as when to encode the picture. While buffered, preprocessing of the input stream by statistics gathering and analysis 310 is performed in accordance with the invention. Pre-encode stage 310 gathers and analyzes statistics on each frame of the sequence of video frames to determine whether the frame has high complexity indicative of noisy video and places the below-described statistics into a stack 314. Stacking of input picture statistics is needed because the GOP structure employed in MPEG encoding of a sequence of video frames may have to be reordered prior to encoding.

When a given frame is to be encoded,
preprocessing 310 thus analyzes the frame to
determine whether one or more encoding parameters
should be adjusted on a macroblock level. As
5 described further below, adjustable parameters may
include macroblock coding type and macroblock
quantization level. This information is forwarded to
the encoder engine 320 commensurate with retrieval of
the frame to be compressed from memory 330. Unless
10 otherwise stated herein, encode engine 320 can
comprise conventional MPEG compression processing as
summarized initially herein.

By way of example, statistics analysis 310
determines whether the current frame has high
15 complexity by determining a statistic equal to an
accumulation of the absolute values of differences
between pairs of adjacent pixels in the frame. This
accumulation is referred to herein as "PIX-DIFF".
PIX-DIFF can be determined by imagining, for example,
20 the luminance data lines of the current picture
concatenated to form a long line of luminance
samples. Then for a given picture, the equation for
the PIX-DIFF statistic might be:

$$25 \quad \text{PIX-DIFF} = \sum_{y=1,3,5\dots}^{\text{Max}} |L_y - L_{y+1}|$$

Where: y is the pixel position number from "1" to the
maximum number of pixels in the concatenated string
of pixels. The PIX-DIFF statistic essentially
30 comprises finding the difference between two adjacent
luminance pixels in this concatenated string of

luminance data for the frame and then summing the
absolute values of those differences. As an
alternative, PIX-DIFF could be defined as an
accumulation of both luminance and chrominance data
5 for the current frame, or an accumulation of
chrominance data only.

Fig. 7 depicts one embodiment for statistics
gathering and analysis in accordance with this
invention. Upon an input picture being available
10 500, statistics processing calculates picture
complexity 510 by determining a PIX-DIFF value for
the picture. A picture with a noisy portion of
random detail will have a very high PIX-DIFF value,
and thus high complexity. The calculated complexity
15 or PIX-DIFF is compared against an experimentally
determined, predefined complexity threshold (TH 1)
520.

Applicants have discovered that in measuring the
PIX-DIFF value for a normal video portion and
20 comparing it to video having a noisy portion, the
noisy portion has a significantly higher PIX-DIFF
value. Thus, if the PIX-DIFF for the frame is less
than the predefined threshold, a noisy picture flag
is set to "0" 530, meaning that the picture comprises
25 normal video only. However, if the complexity of the
picture is high (meaning that the frame contains a
noisy portion), then the target bitrate for the
picture is examined. When the bitrate is high (for
example, 50 Mbits), there may be sufficient bits to
30 encode even a picture with normal and noisy video
portions. Conversely, if the bitrate for the frame
is low, e.g., 4 Mbits, then there may be insufficient

bits to adequately encode the frame. Under this scenario, the encoding options are preferably modified in accordance with this invention. Thus, when the bitrate for the frame is greater than a predefined bitrate threshold (TH 2), the noisy picture flag is set to "0" 530, and when the bitrate is less than this threshold, the noisy picture flag is set to "1" 550. The processing of **Fig. 7** thus results in the setting of a "noisy picture" flag to either "0" or "1". In one embodiment, this flag can be within the statistics analysis 310 preprocessing (**Fig. 6**) and is accessible by the encoder engine 320 upon commencement of encoding of the current frame.

Fig. 8 presents one embodiment for adapting encoding of a picture having a noisy video portion in accordance with the present invention. Picture encoding 600 begins by checking whether the noisy picture flag (**Fig. 7**) has been set 610. If the noisy picture flag is "0", then normal picture encoding 620 is employed. Upon completion of normal picture coding, the encode engine returns 630 to encode the next picture in a sequence of pictures.

On the other hand, if the noisy picture flag has been set, then the macroblock counter is set to "1" 640 and an activity level for each block in the first macroblock is determined 650. The four blocks of the macroblock are ordered based upon their activity level from minimum to maximum and an average block activity is determined from the four values.

If two times the minimum activity level of the blocks is less than the activity level of the next to

minimum block in the macroblock, and two times the minimum activity level in the macroblock is less than the average activity level of the blocks in the macroblock, then the macroblock activity is set to a value equal to the activity level of the next to minimum block in the macroblock. Otherwise, the macroblock activity is set to the minimum activity level in the macroblock 660.

Once the macroblock activity level is set, it is compared against a predefined activity threshold (TH 3) 670. If macroblock activity is below the threshold, then normal macroblock coding 680 is performed; and processing determines whether the macroblock count is at the maximum for the picture 720. If not, the macroblock count is incremented 730 and the activity level for the next macroblock in the picture is calculated. Otherwise, encode processing has been completed, and return is made to process a next picture in the sequence 740.

If the macroblock activity level is greater than the predefined activity threshold (TH 3), then motion estimation is performed 690 and the prediction error or macroblock difference (MBD) is evaluated. If the MBD for the macroblock is greater than, for example, 4096 (4k) and $2 \times (\text{MBD})$ is greater than the macroblock activity level, then the macroblock is coded as an intra (I) macroblock 700. Otherwise, the macroblock is coded as predictive. Once the coding type is determined, the quantization level is calculated 700. The adjusted quantization level is preferably defined as:

ADJ QL=MIN((1 + 0.25 (TH2 - BR + 1)) · CAL QL, MAX ALLOWED BY
STANDARD)

Where: BR is the target bitrate for the
macroblock;

- 5 TH2 is a predefined bitrate threshold;
CAL QL is the calculated quantization level
for the macroblock; and
MAX ALLOWED BY STANDARD is the maximum
quantization allowed by MPEG standard.

- 10 Essentially, the quantization level is increased in
order to conserve bits when the macroblock has high
activity. Once the quantization level is determined,
it is employed in encoding the macroblock. The
macroblock count is then evaluated to determine
15 whether all macroblocks in the picture have been
encoded, and processing continues as described above.

- Those skilled in the art will note from the
description provided herein that processing in
accordance with the present invention prevents noisy
20 macroblocks or blocks with random details from
consuming all or most of the picture bits, which in
turn prevents overproduction of bits before the
encoder reaches the bottom of the picture. This
invention essentially directs encoding bits from the
25 random, busy macroblocks to the simpler, normal
macroblocks. Less bits are used in the highly active
and fine detailed area, and thereby a more constant
picture quality is obtained.

- The present invention can be included, for
30 example, in an article of manufacture (e.g., one or

more computer program products) having, for instance, computer usable media. This media has embodied therein, for instance, computer readable program code means for providing and facilitating the capabilities of the present invention. The articles manufactured can be included as part of the computer system or sold separately.

The flow diagrams depicted herein are provided by way of example. There may be variations to these diagrams or the steps or operations described herein without departing from the spirit of the invention. For instance, in certain cases the steps may be performed in differing order, or steps may be added, deleted or modified. All these variations are considered to comprise part of the present invention as recited in the appended claims.

While the invention has been described in detail herein in accordance with certain preferred embodiments thereof, many modifications and changes therein may be affected by those skilled in the art. Accordingly, it is intended by the appended claims to cover all such modifications and changes as fall within the true spirit and scope of the invention.